

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 15:52:28 ; Search time 170.72 Seconds  
(without alignments)  
17.898 Million cell updates/sec

Title: US-09-331-631A-3\_COPY\_29\_73

Perfect score: 252  
Sequence: 1 SEFROEXECCRQCMOLET.....RCVSQCDKRFEDIDMSKYD 45

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 195891 segs, 67900655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: PIR\_66:\*\*\*  
2: pir1:\*\*\*  
3: pir3:\*\*\*  
4: pir4:\*\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	77.5	30.8	509	2 S08059	alpha-globulin typ
2	77.5	30.8	588	1 FMCNAB	alpha-globulin B p
3	72	28.6	605	2 S06398	alpha-globulin typ
4	70.5	28.0	566	2 S22477	vicilin precursor
5	67.5	26.6	47	2 JC3557	arginine/glutamate
6	67	26.6	810	2 T44430	protein PV100 [imp
7	67	26.6	1170	1 TSHUP1	thrombospondin 1 p
8	67	26.6	1170	2 A40558	thrombospondin 1 p
9	64	25.4	554	2 T45840	GTPase activating
10	63	25.0	524	2 J01730	62k sucrose-bindin
11	59.5	23.6	3078	2 T28432	variant-specific s
12	58	23.0	242	2 T29699	hypothetical prote
13	58	23.0	1421	2 T05892	hypothetical prote
14	57	22.6	316	2 G71600	refin PFB105w - m
15	57	22.6	725	1 MMV294	A-type inclusion p
16	57	22.6	726	2 JQ2162	A-type inclusion p
17	57	22.6	1284	1 MWVZAI	A-type inclusion p
18	55	21.8	308	2 B55346	phosphoprotein pho
19	55	21.8	345	2 A71601	refin PFB1015w - m
20	55	21.8	623	2 T06674	hypothetical prote
21	55	21.8	795	1 HHC008	heat shock protei
22	55	21.8	822	2 T02824	hypothetical prote
23	54.5	21.6	244	2 S44822	P44E2.3 protein -
24	54.5	21.6	1021	2 T15765	hypothetical prote
25	54.5	21.6	1259	2 T13201	hypothetical prote
26	54.5	21.6	2715	2 T13049	eyelid - fruit fly
27	54	21.4	291	2 S62730	cyclin D1 - zebra
28	54	21.4	314	2 T02964	cyclin A-type (cto
29	54	21.4	643	2 T19225	Ro autoantigen 60k

30	54	21.4	880	2 F75103	probable purine NT
31	53.5	21.2	67	2 T15592	hypothetical prote
32	53.5	21.2	301	2 I54209	hypothetical prote
33	53.5	21.2	1642	2 T08880	NMDA receptor-bind
34	53	21.0	622	2 E69006	glutamate synthase
35	53	21.0	702	2 S46854	A28l protein - var
36	53	21.0	702	2 B72167	A29l protein - var
37	53	21.0	702	2 T28570	108k heat shock pr
38	53	21.0	795	2 I50255	ubiquinol--cytochr
39	52.5	20.8	91	1 S00219	collipase A precurs
40	52.5	20.8	96	1 XLH0A	probable prephenat
41	52.5	20.8	275	2 B81430	hypothetical prote
42	52.5	20.8	425	2 T18592	hypothetical prote
43	52.5	20.8	600	2 T18593	thrombospondin 2 p
44	52.5	20.8	1172	1 TSHUP2	antimicrobial pept
45	52	20.6	33	2 A41822	

## ALIGNMENTS

## RESULT 1

S08059  
alpha-globulin type B precursor (tandem 1) - upland cotton (fragment)

N:Alternate names: seed storage protein

C:Species: Gossypium hirsutum (upland cotton)

C>Date: 31-Mar-1990 #sequence\_revision 31-Mar-1990 #text\_change 30-Sep-1993

C:Accession: S08059

R:Chlan, C.A.; Borroto, K.; Kamalay, J.A.; Dure III, L.

Plant Mol. Biol. 9, 533-546, 1987

A:Title: Developmental biochemistry of cottonseed embryogenesis and germination. XIX.

A:Reference number: S06398

A:Accession: S08059

A>Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-509 <CHL>

C:Superfamily: glycinn

Query Match 30.8%; Score 77.5; DB 2; Length 509;  
Best Local Similarity 41.2%; Pred. No. 0.11;

Matches 14; Conservative 9; Mismatches 10; Indels 1; Gaps 1;

OY 5 RQRYEECKRQCMOLETSGQWRRCVSQCDKRFED 38

DB 3 QRRYEECKRQCMOLETSGQWRRCVSQCDKRFED 35

## RESULT 2

FMCNAB  
alpha-globulin B precursor (clone C72) - upland cotton

N:Alternate names: seed storage protein; vicilin precursor

C:Species: Gossypium hirsutum (upland cotton)

C>Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 16-Jul-1999

C:Accession: A30838; S06911

R:Chlan, C.A.; Pyle, J.B.; Legocki, A.B.; Dure III, L.

Plant Mol. Biol. 7, 475-489, 1986

A:Title: Developmental biochemistry of cottonseed embryogenesis and germination XVIII

A:Reference number: A30838

A:Accession: A30838

A:Molecule type: mRNA

A:Residues: 1-588 <CHL>

A:Cross-references: GB:M16891; NID:q167374; PIDN:AAA33071.1; PID:q167375

A:Experimental source: var. Coker 201

R:Chlan, C.A.; Borroto, K.; Kamalay, J.A.; Dure III, L.

Plant Mol. Biol. 9, 533-546, 1987

A:Title: Developmental biochemistry of cottonseed embryogenesis and germination. XIX.

A:Reference number: S06398

A:Accession: S06911

A>Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-81 <CH2>

C:Comment: This is a seed storage protein.



A:Cross-references: GB:X04665; NID:937137; PIDN:CAA28370.1; PID:937138  
A:Note: parts of this sequence, including the amino end of the mature protein, were detected  
R:Labery, C.D.; German, T.M.; Ditty, V.M.  
J. Biol. Chem. 264, 11222-11227, 1989  
A>Title: Characterization of the promoter region of the human thrombospondin gene. DNA s  
A:Reference number: A34274; MUID:89291870  
A:Accession: A34274  
A:Molecule type: DNA  
A:Residues: 1-166 <LAW>  
A:Cross-references: GB:X04835  
R:Hennessy, S.W.; Frazier, B.A.; Kim, D.D.; Deckwerth, T.L.; Baumgartel, D.M.; Rotwein,  
J. Cell Biol. 108, 729-736, 1989  
A>Title: Complete thrombospondin mRNA sequence includes potential regulatory sites in th  
A:Reference number: A30140; MUID:89139590  
A:Accession: A30140  
A:Molecule type: mRNA  
A:Residues: 1-83, 'A', 85-522, 'A', 524-1170 <HEND>  
A:Cross-references: EMBL:X14787; NID:937464; PIDN:CAA32889.1; PID:937465  
A:Note: parts of this sequence, including the amino end of the mature protein, were dete  
R:Kobayashi, S.; Eden-McCutchan, F.; Framson, P.; Bornstein, P.  
Biochemistry 25, 8418-8425, 1986  
A>Title: Partial amino acid sequence of human thrombospondin as determined by analysis c  
A:Reference number: A25812; MUID:87157592  
A:Accession: A25812  
A:Molecule type: mRNA  
A:Residues: 1-83, 'A', 85-397 <KOB>  
A:Cross-references: GB:M25631; NID:9538353; PIDN:AAA36741.1; PID:9538354  
R:Idixit, V.M.; Hennessy, S.W.; Grant, G.A.; Rotwein, P.; Frazier, W.A.  
Proc. Natl. Acad. Sci. U.S.A. 83, 5449-5453, 1986  
A:Reference number: A05172; MUID:86287276  
A:Accession: A05172  
A:Molecule type: mRNA  
A:Residues: 1-83, 'A', 85-374, 'RC', <DIX>  
A:Cross-references: GB:M4436; NID:9340005; PIDN:AAA61237.1; PID:9553801  
A:Note: parts of this sequence, including the amino end of the mature protein, were dete  
R:Sun, X.; Skorstengaard, K.; Mosher, D.F.  
J. Cell Biol. 118, 693-701, 1992  
A>Title: Disulfides modulate RGD-inhibitable cell adhesive activity of thrombospondin.  
A:Reference number: A42927; MUID:92348511  
A:Accession: A42927  
A:Molecule type: protein  
A:Residues: 987-1003 <SDN>  
A:Note: Cys-992 is shown to have a free sulfhydryl  
A:Genetics:  
A:Gene: GDB:THBS1; TSP1; TSP  
F:317-375/Domain: von Willebrand factor type C repeat homology <MAN>  
A:Cross-references: GDB:I20438; OMIM:188060  
A:Map position: 15q15-15q15  
A:Introns: 23/1  
A:Note: the list of introns may be incomplete  
F:Complex: homotrimer, disulfide linked  
F:Function:  
C:Description: participates in cell migration and adhesion, and in platelet aggregation  
C:Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; von  
C:Keywords: beta-hydroxyasparagine; calcium binding; cell adhesion; glycoprotein; trimer  
F:1-18/Domain: signal sequence #status predicted <SIG>  
F:19-1170/Product: thrombospondin 1 #status predicted <MAN>  
F:317-375/Domain: von Willebrand factor type C repeat homology <WMC>  
F:378-499/Domain: thrombospondin type 1 repeat homology <THR1>  
F:434-490/Domain: thrombospondin type 1 repeat homology <THR2>  
F:491-547/Domain: thrombospondin type 1 repeat homology <THR3>  
F:551-586/Domain: EGF homology <EGF1>  
F:650-689/Domain: EGF homology <EGF>  
F:926-928/Region: cell attachment (R-G-D) motif  
F:171-232/Disulfide bonds: #status predicted  
F:248,360,708,1067/Binding site: carboxydrate (Asn) (covalent) #status predicted  
F:620,274/Disulfide bonds: interchain #status predicted  
F:610/Modified site: carboxy-beta-hydroxyasparagine (Asn) #status predicted  
F:1051/Binding site: carboxydrate (Asn) (covalent) #status absent

Query Match	26.6%;	Score 67;	DB 1;	Length 1170;
Best Local Similarity	39.5%;	Pred. No. 3.9;		
Matches	15;	Conservative	6;	Mismatches 11;
				Indels 6;
				Gaps 2

```

QY      13  RQCMQLE-----TSGQMRRC-VSQCDKREEDIDMSKY.44
Db      404  RSCDSLNRCGGSSVQMTCTHQICDCKRKKQDGWSHW 4411

```

RESULT 8  
A40558  
thrombospondin 1 precursor - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 05-Jun-1992 #sequence\_revision 05-Jun-1992 #tex\_change 20-Aug-1999  
C:Accession: A40558; A37905; B42587; S68787  
R:Lewley, J.; Duquette, M.; Ferro, P.; Copeland, N.G.; Gilbert, D.J.; Jenkins, N.A.  
Genomics 11, 587-600, 1991  
A:Title: Characterization of the murine thrombospondin gene.  
A:Reference number: A40558; MUID:92128941  
A:Accession: A40558

A:Accession: A40306  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1170 <LAW>  
A:Cross-references: GB:M62449; GB:M62450; GB:M62451; GB:M62452; GB:M62453; GB:M62454; GB:M62455; GB:M62456; GB:M62457; GB:M62468; GB:M62469; GB:M62470; NID:9511867; PIDN:AAAS  
R:Bornstein, P.; Alfi, D.; Devarayalu, S.; Franson, P.; Li, P.  
J. Biol. Chem. 265, 16691-16698, 1990  
A:Title: Characterization of the mouse thrombospondin gene and evaluation of the role  
A:Reference number: A37505; MUID:50375546  
A:Accession: A37505

A:Accession: A57905  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-490 <B0R>  
 A:Cross-references: GB:J05605; GB:J05606; NID:g201991; PIDN:AAA4031.1; PID:g554390  
 L:Alaery, C.D.; O'Rourke, K.; Wolf, F.W.; Katz, R.; Seldin, M.F.; Dixit, V.M.  
 I: Biol. Chem. 267 3274-3281 1992

U. Biol. Chem. 261: 9214-9217, 1986  
A>Title: Characterization of mouse thrombospondin 2 sequence and expression during ce  
A:Reference number: A42587; MUID:92147663

A/Accession: B42587  
A/Status: preliminary; not compared with conceptual translation  
A/Molecule type: mRNA

A;Residues: 1-1152, 'P', 1154-1170 <LAH>  
A;Cross-references: GB:M87276  
A;Note: sequence extracted from NCBI backbone (NCBIP:81501,

**A:**Title: Expression and initial characterization of recombinant mouse thrombospondin A;Reference number: S68787; MUID:96234006

A:Accession: S68/U/  
A:Molecule type: protein  
A:Residues: 19-26,'X',28-37 <CHE>  
C:Complex: homotrimer, disulfide linked  
C:Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology;  
;Keywords: calcium binding; glycoprotein; homotrimer

Accession	Protein	Signal	Sequence	#status	predicted	<Sig>	<AMT>	<VNC>
F_1-18/Domain	thrombospondin	1	thrombospondin	1	status	predicted	<Sig>	<AMT>
F_15-1170/Product	thrombospondin	1	thrombospondin	1	status	predicted	<AMT>	<VNC>
F_317-375/Domain	von Willebrand factor	type C	repeat	homology				
F_378-429/Domain	thrombospondin	type 1	repeat	homology				<THN1>
F_434-490/Domain	thrombospondin	type 1	repeat	homology				<THN2>
F_431-547/Domain	thrombospondin	type 1	repeat	homology				<THN3>
F_551-586/Domain	EGF	homology	<EGF>					
F_248-360/708,1067	banding site	carbohydrate	(Asn)	(covalent)	#status	predicted		

		Best Query Match	26.6%	Score 67;	DB 2;	Length 110;
		Pest Local Similarity	39.5%	Pred. No. 3.9;		
		Matches	15;	Conservative	6;	Mismatches 11; Indels 6; Gaps 2.
OY	13	RQCMOLF-----TSGQMRRC--VSQCDCRPFEDIDMSKY	44			
Db	404	RSCDSLNNRCGSSVOTRTCHIQECDKRFKKDGGGSHW	441			

RESULT 9  
T45840  
GTPase activating-like protein - Arabidopsis thaliana

```

N:Alternate names: protein F2K15.210
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
C:Accession: T45840
R:Lieger, M.; Gabell, C.; Mueller-Auer, S.; Schaefer, M.; Zipp, M.; Mewes, H.W.; Lemcke,
submitted to the Protein Sequence Database, January 2000
A:Reference number: 223015
A:Accession: T45840
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-554 <EMB>
A:Cross-references: EMBL:AL32956
C:Genetics:
A:Map position: 3
A:Introns: 53/2; 98/3; 115/2; 133/2
A>Note: F2K15.210

Query Match          25.4%; Score 64; DB 2; Length 554;
Best Local Similarity 28.2%; Pred. No. 4.5;
Matches 11; Conservative 10; Mismatches 18; Indels 0; Gaps 0;

.Oy      5 ROYECKRCQCMQLETSQGMRCVSCDCRFREDIDMSK 43
         ||||| :||| :|: |::| ::||| :
Db       148 RKEVERLRQCKRLQKHNNGTAKLYNGSETTIDETDWR 186

RESULT 10
J01730
62K sucrose-binding protein precursor - soybean
C:Species: Glycine max (soybean)
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 26-Aug-1999
C:Accession: J01730
R:Grimes, H.D.; Overvoorde, P.J.; Ripp, K.; Franceschl, V.R.; Hiltz, W.D.
Plant Cell 4, 1561-1574, 1992
A>Title: A 62-kD sucrose binding protein is expressed and localized in tissues actively
A:Reference number: J01730; MUID:93104680
A:Accession: J01730
A:Molecule type: mRNA
A:Residues: 1-524 <GRI>
A:Cross-references: GB:L06038; NID:g1431744; PIDN:AAB03894.1; PID:g170064
C:Superfamily: glycinin
C:Keywords: sugar transport
F:1-29/Domain: signal sequence #status predicted <SIG>
F:30-524/Product: 62K sucrose-binding protein #status predicted <MAT>

Query Match          25.0%; Score 63; DB 2; Length 524;
Best Local Similarity 38.2%; Pred. No. 5.6;
Matches 13; Conservative 4; Mismatches 15; Indels 2; Gaps 1;

Oy      2 EFDRQEYEECKRCQMQL--TSGQMRRCVSQCDK 33
         |::| |::| |::| |::| |::| |::| |::|
Db       34 EEDEPELVTCRKHCQQQQQQYTGGDKRVCLQSDDR 67

RESULT 11
T28432
variant-specific surface protein 1 - malaria parasite (Plasmodium falciparum)
N:Alternate names: erythrocyte membrane binding protein 1 (EMP1)
C:Species: Plasmodium falciparum
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000
C:Accession: T28432
R:Su, X.Z.; Heatwole, V.M.; Wertheimer, S.P.; Guinet, F.; Herrfeldt, J.A.; Peterson, D.S.;
Cell 82, 89-100, 1995
A>Title: The large diverse gene family var encodes proteins involved in cytoadherence ar
A:Reference number: Z20487; MUID:95330813
A:Accession: T28432
A>Status: preliminary; translated from GR/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-3078 <SUN>
A:Cross-references: EMBL:L40608; NID:g886374; PID:g886375; PIDN:AAA75396.1

```

[illegible]



